J. Dairy Sci. 85:459–471

Molten Globule Structures in Milk Proteins: Implications for Potential New Structure-Function Relationships[1]

H. M. Farrell Jr.,* P. X. Qi, E. M. Brown, P. H. Cooke, M. H. Tunick, E. D. Wickham, and J. J. Unruh
U.S. Department of Agriculture, Agricultural Research Service[2]
Eastern Regional Research Center, 600 E. Mermaid Lane
Wyndmoor, PA 19038

ABSTRACT

Recent advances in the field of protein chemistry have significantly enhanced our understanding of the possible intermediates that may occur during protein folding and unfolding. In particular, studies on α-lactalbumin have led to the theory that the molten globule state may be a possible intermediate in the folding of many proteins. The molten globule state is characterized by a somewhat compact structure, a higher degree of hydration and side chain flexibility, a significant amount of native secondary structure but little tertiary folds, and the ability to react with chaperones. Purified α_{s1}- and κ-caseins share many of these same properties; these caseins may thus occur naturally in a molten globule-like state with defined, persistent structures. The caseins appear to have defined secondary structures and to proceed to quaternary structures without tertiary folds. This process may be explained, in part, by comparison with the architectural concepts of tensegrity. By taking advantage of this "new view" of protein folding, and applying these concepts to dairy proteins, it may be possible to generate new and useful forms of proteins for the food ingredient market.
(**Key words:** casein structure, whey proteins, protein folding, spectroscopy of proteins)

Abbreviation key: AFU = autonomously folded units, **κ-CN** = κ-casein, **CD** = circular dichroism, **DSC** = differential scanning calorimetry, **EM** = electron microscopy, **FTIR** = Fourier transform infra-red, **α-LA** = α-lactalbumin, **MD** = molecular dynamics, **PIPES** = (piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)), **RCM** = reduced carboxymethylated, **SAXS** = small-angle X-ray scattering.

INTRODUCTION

The recent announcement that the human genome has been solved represents not an end in itself, but a beginning. It is now the work of the protein chemists and structural biologists to translate this linear sequence information into the functional molecular shapes necessary to make significant breakthroughs in biological applications. This task is already under way and, by the union of experimental and theoretical protein studies, we may soon be able to more fully understand how protein structure produces biological function. Basic studies on the protein-folding problem are already producing a wealth of fundamental information (Arai and Kuwajima, 2000; Jaenicke and Lilie, 2000; Onuchic et al., 2000; Peng and Wu, 2000). These basic studies can be translated directly into food applications. It has long been known that changing protein structure can alter food functionality. We have often previously thought of these changes as all (completely unfolded protein) or nothing (completely native protein) events. The latest studies in protein folding, however, demonstrate a multiplicity of stages in protein folding (unfolding) pathways. The energy profiles suggest that these "molten globule" states may be stable under certain conditions (Arai and Kuwajima, 2000). If this is the case, then a single whey protein such as α-lactalbumin may have a myriad of stable intermediate stages that could be "trapped" and thus provide the potential for new functional properties. As food protein chemists we must also recall that while native structure arises from sequence, processing treatment has the potential to transform native structure into nonnative states. This latter fact represents both a challenge and an opportunity, as denatured, or partially denatured, proteins may serve as either a problem or a potentially valuable food ingredient. Examples of these opportunities are given for α-lactalbumin (**α-LA**) and κ-casein (**κ-CN**).

Received October 9, 2000.
Accepted February 15, 2001.
Corresponding author: H. M. Farrell, Jr.; e-mail: hfarrell@arserrc.gov.
[1]Presented in the Symposium "Biologically Active Peptides" sponsored by the Milk Protein and Enzyme Nomenclature Committee at the ADSA, ASAS Joint Meeting, Baltimore, MD, July, 2000.
[2]Mention of brand or firm name does not constitute an endorsement by the U.S. Department of Agriculture above others of a similar nature not mentioned.

MATERIALS AND METHODS

Materials

All reagents used were of analytical grade or 'ACS certified' from Sigma (St. Louis, MO). Isolation and purification of κ-CN, its compositional analysis, reduction, and alkylation were presented in some detail elsewhere (Farrell et al., 1996, 1999; Groves et al., 1992). α-LA was purified as previously described (Akers et al., 1986).

Circular Dichroism (CD) Measurements

Far-UV CD experiments were carried out with 0.3 mg/ml protein samples in a 2 m*M* piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (**PIPES**), 4 m*M* KCl buffer at pH 6.75. Successive measurements in the far UV (190 to 250 nm) were made at 10, 25, 50, and 70°C. Solvents for CD measurements were first filtered through a Millipore 0.22-μm pore filter. Dissolved protein samples were filtered through centrifuge tubes containing a 0.45-μm pore regenerated cellulose filter. The CD spectra were recorded on an Aviv model 60DS spectropolarimeter (Aviv Associates, Lakewood, NJ), using cells of appropriate path lengths and a scan time of 40 s/nm. The jacketed cells were attached to a circulating constant temperature bath; the time for equilibration of the sample was calculated to be 30 min for a 10° change in the bath temperature. Near UV spectra of α-LA were conducted as previously described (Alaimo et al., 1999). Spectra are corrected for solvent contributions and are expressed in units of mean residue ellipticity (far-UV) or molar ellipticity (near-UV) versus wavelength. Analysis of peptide secondary structure from CD spectra was accomplished as previously described (Alaimo et al., 1999).

Molecular Dynamics Simulations

α-Lactalbumin denaturation was simulated in vacuo (PDB code, 1HFZ) from Pike et al. (1996). Molecular dynamics (**MD**) simulations were carried out with the Tripos Sybyl molecular modeling software (Version 6.61) on a Silicon Graphics Indigo 2 workstation. The Tripos force field and a Kollman all-atom charge set were used throughout the simulations. A 20-step, conjugated-gradient energy-minimization was employed for all structures after each dynamics simulation. A cutoff value of 8 Å was used for all nonbonded interactions.

Infrared (FTIR) Spectroscopic Measurements

For infrared (**FTIR**) measurements, 3 mg of κ-CN was dissolved as a 3.0% wt/wt aqueous solution at pH 6.75 in 25 m*M* dipotassium PIPES buffer containing 80 m*M* KCl. Spectra were obtained at 25°C using a Nicolet 740 FTIR spectrometer (Madison, WI) equipped with a Nicolet 660 data system. Data collection was carried out after a nitrogen purge of the sample chamber, which consisted of a demountable cell with CaF_2 windows separated by a 6-μm Mylar spacer. Each spectrum consists of 4096 double-sided interferograms, coadded, phase-corrected, apodized with a Happ-Genzel function, and fast-Fourier transformed. Secondary structural features were calculated from the amide I envelope, after assignment of putative peaks by use of second derivative spectra, and fitting of Gaussian peaks to the original spectra. The selected peaks were refitted to the original spectra using an iterative curve-fitting procedure as previously reported (Kumosinski and Unruh, 1996). Data are reported as the sum of assigned spectral elements gathered from a 20-protein database (Kumosinski and Unruh, 1996).

Electron Microscopy of κ-CN

Samples of the κ-CN were prepared for electron microscopy by dissolving the casein in 25 m*M* PIPES buffer pH 6.75 made to be 80 m*M* KCl. The samples were made up to be 30 to 35 mg/ml and were passed through 0.45-μm filters. The filtrates were adjusted to 25 mg/ml with filtered buffer and equilibrated at 25 or 37°C for 30 min.

Thin support films of amorphous carbon were evaporated on strips of cleaved mica and mounted on 400 mesh copper grids. All subsequent procedures were carried out over a water bath with samples and reagents at 37°C. Aliquots (10 μl) of casein in buffered solution were placed on freshly prepared support films for 30 to 60 s over a water bath at 37°C; then the sample-side of the grid was washed with a controlled stream of 10 to 15 drops of buffered solution from a disposable Pasteur pipette containing 1% glutaraldehyde at 37°C. This was done to physically stabilize the composition of monomers in the form of polymers and to trap the equilibrium structures, while reducing the protein concentration to produce a discontinuous monolayer of κ-CN particles. Then the adsorbed particles were washed with a similar controlled stream of 5 to 10 drops of 2% uranyl acetate solution at 37°C for negative staining. Excess uranyl acetate solution was adsorbed from the grid surface into Whatman #1 filter paper, and grids were allowed to air dry at room temperature.

Images of κ-CN structures in randomly selected fields on grids were recorded photographically, at instrumental magnifications of 88,000× with a Zeiss model 10 B electron microscope (Thornwood, NY) operating at 80KV, and 45,000 or 97,000× using a Philips model CM12 scanning-transmission electron microscope

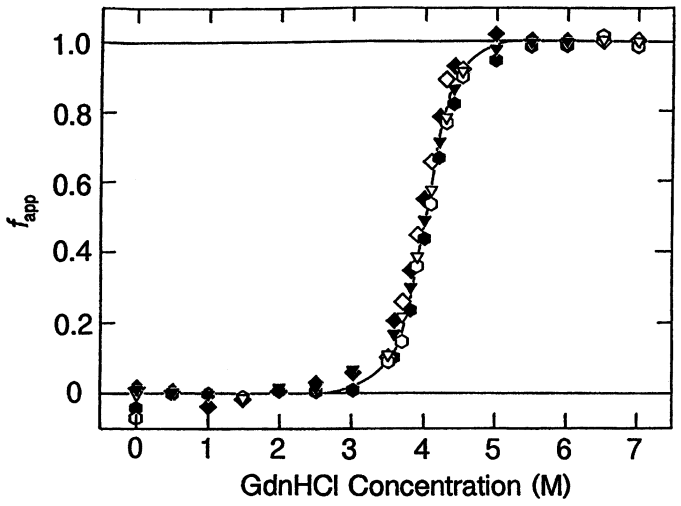

Figure 1. Guanidine HCl induced unfolding of lysozyme showing coincident loss of enzyme activity and of structure as measured by f_{app}, the apparent fraction denatured. Open symbols represent NaCl only and closed represent NaCl + $CaCl_2$. Changes in circular dichroism were monitored at 289 nm (◆, ◇) 255 nm (filled hexagon, open hexagon) and 222 nn (▼, ▽). Data from Arai and Kwajima (2000) reproduced courtesy of Academic Press, San Diego, CA.

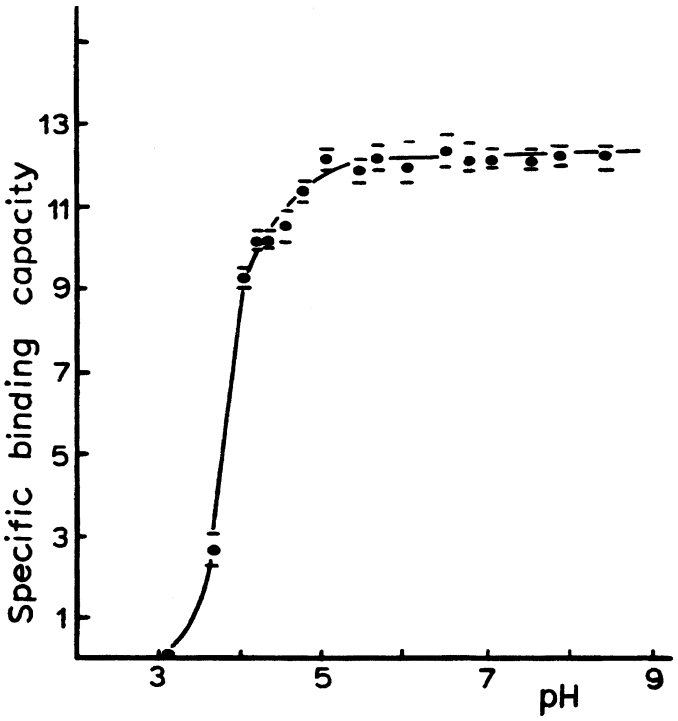

Figure 2. The pH induced loss of riboflavin binding of riboflavin binding protein, with parallel loss of protein structure; data from Kumonsinski et al. (1982).

(Hillsboro, OR) operating at 60 KV in the bright field imaging mode.

RESULTS AND DISCUSSION

Historical Background

Historically, the central dogma of structural biology is the Anfinsen hypothesis: the linear primary sequence of amino acids of a protein codes for rather specific secondary structural elements, which in turn lead to protein folding and tertiary structure, and ultimately to quaternary structure for complex higher order systems (Arai and Kuwajima, 2000; Jaenicke and Lilie, 2000; Onuchic et al., 2000; Peng and Wu, 2000). There has been considerable debate in recent years, not so much on the veracity of this hypothesis, but on the details of the mechanisms that bring about protein structure. This debate arises, in part, because until very recently our best information on protein folding came from studies of protein unfolding. The early studies on lysozyme, as summarized by Arai and Kuwajima (2000), indicated a concerted mechanism by which loss of enzyme activity and loss of structure followed nearly similar curves. Figure 1 shows the typical sigmoidal curve found for lysozyme denaturation by guanidine HCl, showing loss of structure; this was coincident with loss of enzymatic function (Arai and Kuwajima, 2000). A similar curve was constructed for egg white riboflavin binding protein (Figure 2), where loss of structure and ability to bind riboflavin were coincident upon change in pH as shown by Kumosinski et al. (1982). These and many other studies could be summed up in equation 1.

$$\text{Native} \underset{K}{\Leftrightarrow} \text{Denatured} \qquad [1]$$

Here K was considered to represent a cooperative process (based upon the sigmoidal nature of the curves) such as those shown in Figures 1 and 2. By inference, protein folding was thought to be the reverse reaction. Indeed the argument was made that kinetically the two processes might be microscopically reversible, and that the native state could represent a singular folded state with minimum potential energy.

α-Lactalbumin and the Molten Globule State

The "new view" of protein folding began to emerge, in part, from experimental observations on the subject of much of this symposium: α-LA. Physical studies on this protein appeared to indicate that some measurable structural loss might occur at different stages in the denaturation process. Conventional FTIR and far-UV CD spectroscopies are used to follow losses of secondary

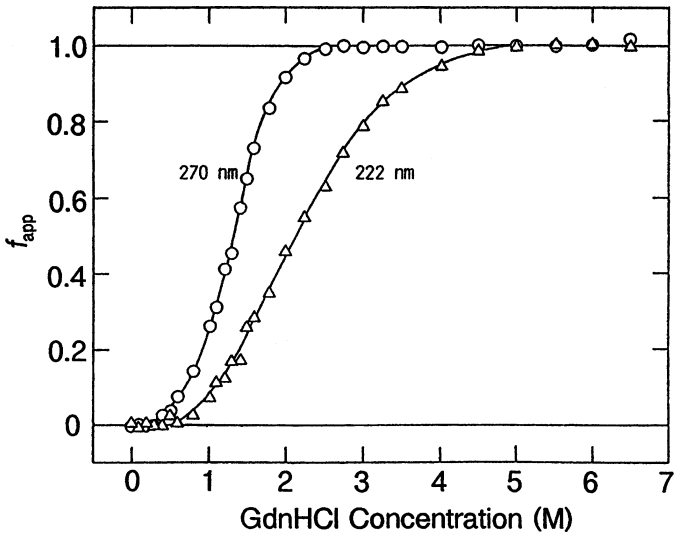

Figure 3. Guanidine HCl induced unfolding of α-lactalbumin; the f_{app} (fraction denatured) shows noncoincident loss of protein structure by circular dichroism. Loss of aromatic dichroism was measured at 270 nm (○); loss of α-heliz was followed at 222 nm (△). Data from Arai and Kwajima (2000) reproduced courtesy of Academic Press, San Diego, CA.

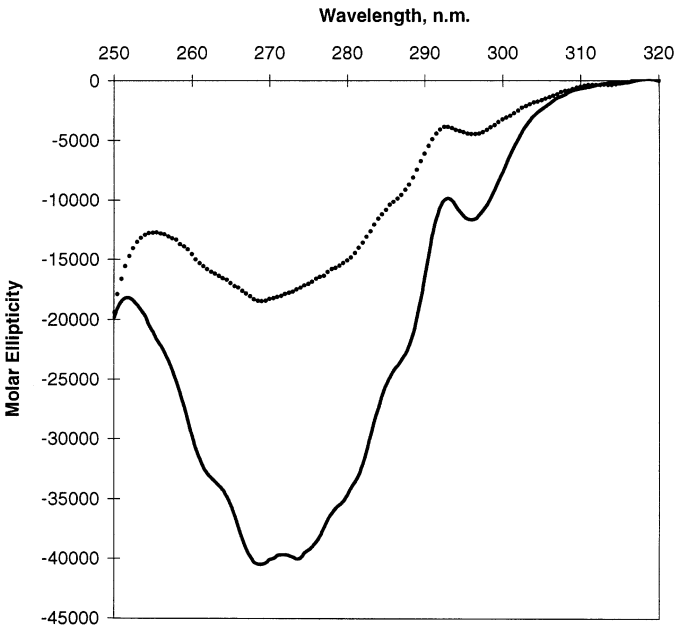

Figure 4. Heat induced loss of aromatic circular dichroism spectra of α-lactalbumin in PIPES buffer at pH 6.75. Lower curve at 25°C, upper curve at 50°C. Molar ellipticity is in degrees cm^2/dmol. The protein concentration was 62.8 μM.

structural elements, but near-UV CD can also provide important information on protein tertiary structure. In a typical globular protein, steric restrictions as well as hydrogen bonding can limit the mobility of aromatic side chains such as Trp, Tyr, and Phe, thus causing a near-UV CD spectra. Unrestricted rotation of aromatic groups does not in general give rise to a CD spectrum. Figure 3 shows that for guanidine denaturation of α-LA, loss of aromatic dichroism preceded the loss of secondary structure. Figure 4 shows the typical near-UV CD spectra of bovine α-LA, at 25 and 50°C. A significant amount of the near-UV CD spectrum has degraded by the time the temperature has reached 50°C. In contrast, measures of secondary structural change reveal higher transition temperatures. Differential scanning calorimetry (**DSC**) studies of α-LA by Dolgikh et al. (1985) showed that the temperature of denaturation is about 60°C (Figure 5). When the bound Ca^{2+} is released from α-LA there is a loss of tertiary structural stability and the temperature of denaturation lowers to about 30°C (Figure 5). Curiously, on acid denaturation at pH 2.0, much secondary structure is retained, but a complete loss of the characteristic thermal transition of α-LA occurs (flat line, Figure 5). The lack of coincidence of loss of aromatic CD and secondary structure began the search for a broader meaning of Equation 1.

Friere and his colleagues (Xie et al., 1991) introduced the concept of three-dimensional DSC by which the temperature of denaturation was studied as a function of guanidine concentration. Surprisingly, 1 *M* guanidine alleviated almost all of the typical enthalpic thermal transition found for α-LA in the absence of guanidine, yielding a DSC pattern similar to that found at pH 2.0 (Figure 5). Recalling the data of Figure 3, it can be seen that at 1 *M* guanidine, there has been some

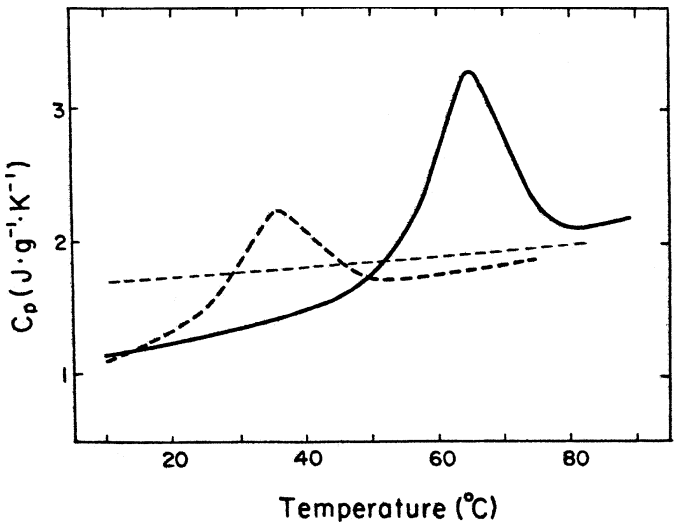

Figure 5. Dependence of partial heat capacity (Cp) on temperature for α-lactalbumin as holoprotein (—) apoprotein (dark dashed line) and acid denatured protein (flat, light dashed line). Data from Dolgikh et al. (1985) reproduced courtesy of Springer-Verlag Heidelberg, Germany.

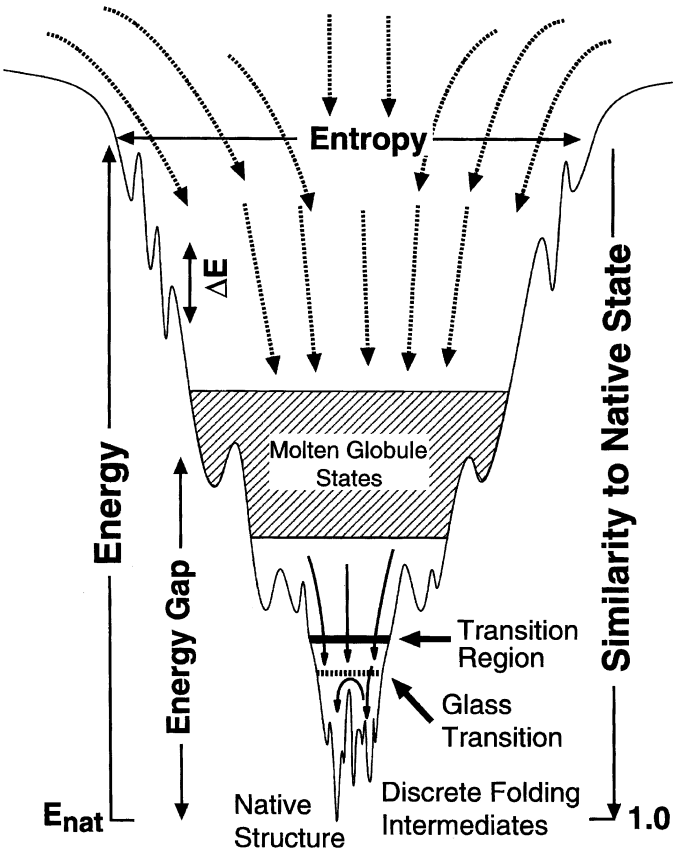

Figure 6. A schematic representation for the energy landscape for a minimally frustrated heteropolymer during protein folding. E_{nat} is the minimum potential energy for the native state. From Onuchic et al. (2000); reprinted courtesy of Academic Press, San Diego, CA.

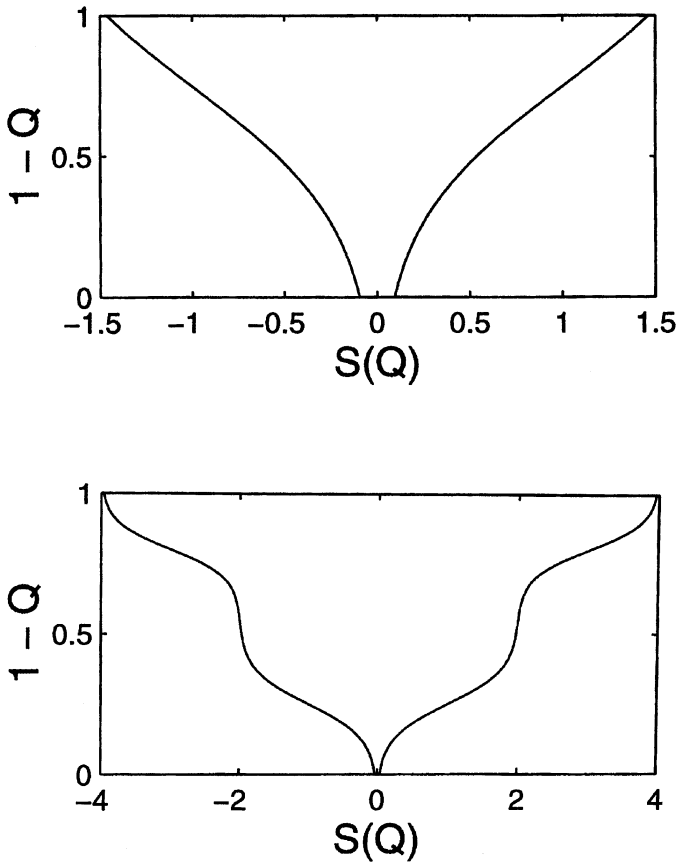

Figure 7. More simplified representations of energy landscapes for a rapidly folding protein (top) and a protein with a variety of closely related intermediates (bottom); the axes relate to Figure 6. The value (1-Q) on the ordinate is equivalent to the right hand ordinate of Figure 6. The abscissa S (Q) is proportional to Entropy depicted at the top of Figure 6. Figure adapted from Onuchic et al. (2000); courtesy of Academic Press, San Diego, CA.

change in side chain dichroism, but little change in secondary structure. Studies on many other proteins now reveal that multi-step mechanisms of protein unfolding may be the rule, rather than the exception. The condition in which a partially denatured protein may exhibit a high degree of segmental motion (i.e., loss of aromatic dichroism) while retaining a significant amount of secondary structure has been termed the molten globule state.

Energy Landscape Theory

The idea that a protein may unfold through a multistep process in turn leads to the "new view" of protein folding. During folding, a protein chain may "sample" a significant amount of conformational space before settling into a selective energy minimum. Indeed, several false minima may lie quite close or even somewhat remotely removed from the true minimum. Such an intermediate area has been postulated to be the molten globule state. This overall idea as presented by Onuchic et al. (2000) can be viewed in two dimensions in Figure 6 for the theoretical energy landscape of a "frustrated" heteropolymer, as folding is viewed from top to bottom. The parameter Q is the global order parameter and represents a value of 1.0 for all interactions in the native state, and 0.0 for the completely unfolded protein. Reflection of the interactions between Q, E, and S result in a three-dimensional funnel. In actuality protein folding funnels may be rather simple (Figure 7, top), complex (Figure 6), or goblet shaped (Figure 7, bottom) with a variety of closely related intermediate states. The important prediction here is that Equation 1 can now be modified:

$$\text{Native} \Leftrightarrow \text{Intermediate} \Leftrightarrow \text{Denatured} \qquad [2]$$

Here the intermediate stage may be represented by a multiplicity of macroscopic states or ensembles as shown in Figure 6. The complexity of this picture offers

an enormous challenge to the study of protein folding. Conversely, a myriad of states in unfolding can represent an enormous opportunity for the food protein chemist, as stable folding intermediates or partially unfolded proteins may offer new food functionality. This is not to say that there are a limitless number of possible denatured conformational states for a protein or that the native state is without mobility or internal motion. The amount of conformational space sampled by a given protein is limited by a number of factors. These include the innate, somewhat planar, nature of the peptide bond, and its energetic preference for the trans conformation, as well as restriction in rotation due to amino acid side chains (Peng and Wu, 2000). In addition, certain combinations of nearest neighbors can influence the propagation of secondary structure. For these reasons, sequence-based secondary structural prediction algorithms work quite well in some cases. In others they fail because tertiary interactions from distal parts of the molecule, or another molecule (chaperones) may cause the protein to accelerate on a fast track to a folded compact structure (Jaenicke and Lilie, 2000). However, autonomously folded units (**AFU**) of proteins appear to be the rule, as discussed by Peng and Wu (2000). Such AFU may exist as stable units within the molten globule state or be liberated from the protein by proteolytic action. The AFU concept also offers the possibility of generating new protein-based products with altered biological or neutraceutical activity or new food functionality.

Hydration and Stability of the Molten Globule State

The typical molten globule state is more open than the native state, but still represents a compact structure. The increased flexibility of the side chains and the backbone leads to an influx of water, so that the molten globule state is rather highly hydrated (Onuchic et al., 2000). This water does not represent an increase in bound water, but rather a type of protein influenced or "trapped but exchangeable" water such as that studied by NMR relaxation techniques (Mora-Gutierrez et al., 1997). In addition, the loosely packed side chains may emphasize the importance of solvent-separated hydrophobic interactions in the molten globule state, whereby a water molecule separating two hydrophobic residues lends a small degree of stabilization to protein structures.

Three-Dimensional Structures and Molecular Modelling Studies

To explore the range of possible states of denaturation available for α-LA, we conducted a series of rapid MD simulations in vacuo. These simulations mimic the action that heat or shear would have on the folded polypeptide chain. In essence, by adding energy to the model, we shift the molecular energy profile from bottom to top in Figure 6. We used the bovine α-LA protein database code 1HFZ, which contains bound Ca^{2+} (holoprotein), and compared it to the same structure with the Ca^{2+} removed (apoprotein). The MD simulations were conducted at 5° and 298°K for the apo and holo forms. The 5°K simulation for the apoprotein in vacuo rapidly came to equilibrium (overnight) as judged by the change of the radius of gyration (Rg) with time (Figure 8). Similar curves were obtained for the other simulations. This equilibrium position implies a possible denatured state with some stability, perhaps in a trap somewhere near the middle of Figure 6. Comparisons of two of the models generated for the apoprotein form of α-LA by the MD simulations with native α-LA are shown in Figure 9. Here it can be seen that removal of Ca^{2+} followed by MD results in two different unfolded states at two different temperatures. The analyses of these models are presented in Table 1, where they are compared with experimental small-angle X-ray (**SAXS**) data for bovine α-LA. The Rg at 5°K increased from 14.3 to 18.4 Å, which is comparable to the overall change found for MD simulations of human α-LA carried out in explicit water for days (Smith et al., 1999). However, in our simulations, secondary structure was rapidly lost, whereas in the human α-LA in water, they were reasonably constant. Our rapid simulations, however, appear to yield Rg values for the holoprotein, which are only 7% larger than the molten globule state, but

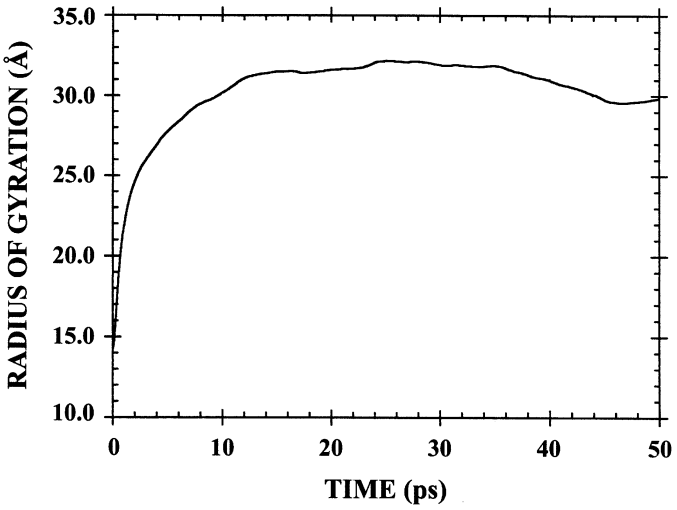

Figure 8. Molecular dynamics simulation of the apoform of α-lactalbumin at 5°K in vacuo. The Rg (radius of gyration) values begin at that of the native, and then rapidly approach an equilibrium point, indicating attainment of a simulated altered state for the protein.

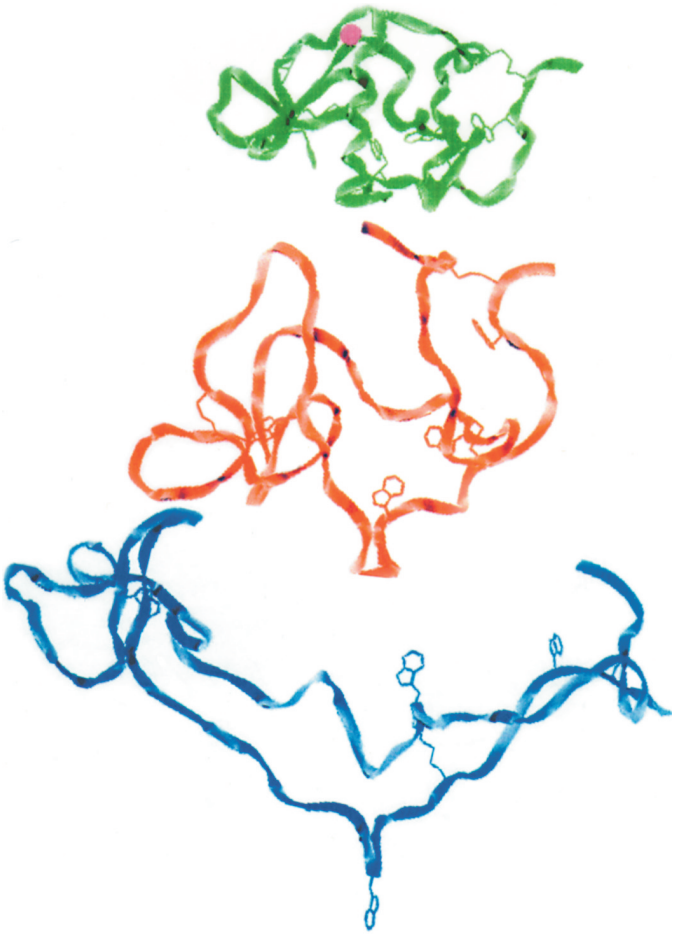

Figure 9. Rapid molecular simualtions of the apoprotein form of α-lactalbumin in vacuo, showing the native holo state (top) and the effect of simulations at 5 and 298°K (middle and bottom, respectively) of the apoform.

smaller than those of the urea denatured structure (Table 1). In the case of the apoprotein, the Rg value at 298°K approximates that of the urea denatured protein. Thus, removal of the Ca^{2+} ion leads to greater destabilization of the molecule. This is in accord with experimental data, which show in DSC a much-lowered temperature of denaturation for the apoprotein (Figure 5). Thus, the rapid MD simulations are quite instructive in that they predict a multiplicity of quasi-stable states for the protein beyond the molten globule state, as well as demonstrate the outer limits for denaturation of the protein with and without Ca^{2+} and in the absence of disulfide rupture.

Table 1. Comparison of molecular dynamics simulated Rg values with experimental data for bovine α-lactalbumin (α-LA).

	Calculated Rg (Å)	SAXS Experimental Rg (Å)[1]
Holo α-LA	14.3	15.7
3.5 ps 5°K	18.4	—
40 ps 298°K	21.9	—
Molten globule	—	17.2
Apo α-LA	14.3	—
1.5 ps 5°K	21.4	—
20 ps 298°K	31.6	—
Urea unfolded	—	30.0

[1]Kataoka et al. (1997). SAXS = small-angle X-ray scattering.

Table 2. Interatomic distances for selected tryptophan residues in bovine and human α-lactalbumin crystal structures.

	Crystal structure	Distances (Å)
Bovine residues[1]	26 to 104	6.8
	26 to 118	12.4
	60 to 104	9.2
Human residues[2]	31 to 103	6.8
	31 to 118	3.4
	60 to 103	3.9

[1]This study.
[2]Saito, 1999 and Smith et al., 1999.

In studies on human α-LA, Smith et al. (1999) have shown that there are two pronounced aromatic clusters in the crystal structure. Cluster I comprises Trp 118, Tyr 36, Phe 31, and Leu 27, while cluster II includes Trp 60, Tyr 103, and Trp 104. These restricted aromatic residues are the source of the near-UV CD spectra. The three-dimensional view of bovine α-LA shows a somewhat different arrangement of aromatic residues. In the bovine molecule, Trp 104 is in closer proximity to Trp 26 (Leu 27 in human) than to Trp 60. Thus, Trp 104 shifts from one cluster to the other in the bovine protein's structure. This difference in tertiary folding is shown schematically in Figure 10 for a linear structure of bovine α-LA. This change in alignment places Trp 118 on the surface of bovine α-LA, and causes Trp 60 to be more exposed as well. Table 2 gives the calculated distances between selected residues in the bovine and human crystal structures. For the human and baboon structures, long-term aqueous MD simulations (Saito, 1999; Smith et al., 1999) show that for cluster I an average change of interatomic distances from 3.8 to 8.2 Å (over twofold) occurs on going from the native to the molten globule state. This indicates a dramatic opening up of aromatic cluster I. Cluster II, however, moves less dramatically (4.4 to 6.5 Å), perhaps due to the strong attractions between Trp 60 and Trp 104. Because the bovine protein lacks this interaction, larger changes associated with unfolding would be expected. Table 3 shows that increases in interatomic distances for the aromatic residues in our rapid simulations range from two- to threefold. Experimentally, this corresponds to the temperature dependent loss of aromatic CD (Figure 4). Comparison of the values in Table 3 with

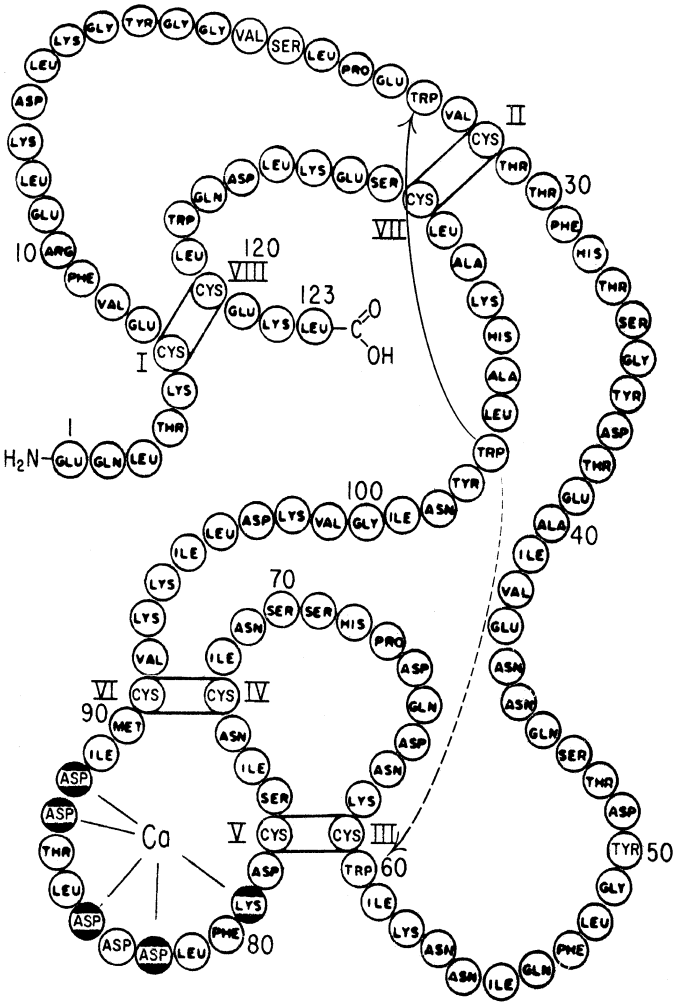

Figure 10. Linear representation of the sequence of bovine α-lactalbumin. The solid arrow denotes the direction of folding for Trp 104 in the bovine protein, while the dotted arrow represents the folding pattern for the human and baboon forms; the Ca^{2+} binding site is indicated by shaded residues.

those of Table 1 indicates that the net movement of the aromatics upon denaturation rivals that of the overall change in Rg. The changes are predicted to be even greater without Ca^{2+}. This indicates that the bovine α-LA would have more new surface aromaticity for protein-protein interactions upon partial denaturation even without disulfide rupture. These new hydrophobic contacts could lead to new potential protein products from whey isolates. Such products could have neutraceutical applications.

Partial Denaturation and Aggregation

From the perspective of the natural processes that occur in protein folding, association is the process whereby structured monomers, which on the general folding pathway, form their multimeric complexes as usually found in native state. In contrast, protein-protein interactions from the molten globule state, may lead to aggregation as defined by Jaenicke and Lilie (2000). The results of MD simulations described above suggest that many new contact areas for protein-protein interactions may be available in a partially denatured state such as the molten globule state or upon further denaturation as demonstrated in Figure 9 for α-LA. Research on human α-LA has demonstrated that polymeric forms of the protein may serve as antitumor and antibacterial agents (Håkansson et al., 2000; Svensson et al., 1999). In a series of experiments designed to purify bovine α-LA for the preparation of highly specific polyclonal antibodies (Akers et al., 1986), we chromatographed the protein on DEAE-cellulose in the presence and absence of Ca^{2+}. The final product was dialyzed, lyophilized, and then dialyzed against Ca^{2+} and placed on a G-100 Sephadex column in imidazole buffer. Two peaks emerged (Figure 11). The peaks were divided into fractions I and II and analyzed by SDS gels. Both fractions were found to contain only α-LA. For the previous study (Akers et al., 1986) only the monomeric-Fraction II was used for the antibodies. This raises the question of whether or not the polymeric α-LA would have given rise to a protein with altered biological activity. Regardless of the biological activity, this experiment demonstrates clearly that partial denaturation of α-LA with and without Ca^{2+} can lead to new and potentially interesting food ingredients. Con-

Table 3. Changes in the interatomic distances in Ångstroms for CH_2 aromatic carbon atoms tryptophan residues before and after molecular dynamics simulations for bovine α-lactalbumin (α-La) in vacuo.

Molecules	Trp60–Trp104[1]	Trp26–Trp104[1]	Trp26–Trp118[1]
α-La X-ray structure[2]	9.17 ± 0.83	6.84 ± 1.93	12.42 ± 1.39
Apo α-La 5°K & 1.5 ps	19.14 ± 1.20	10.68 ± 1.60	18.95 ± 1.66
Apo α-La 298°K & 20 ps	49.12 ± 0.25	33.41 ± 3.13	21.71 ± 0.76
Holo α-La 5°K & 3.5 ps	17.97 ± 1.62	11.89 ± 2.38	19.36 ± 1.66
Holo α-La 298°C & 40 ps	17.32 ± 1.04	23.72 ± 2.89	30.04 ± 1.90

[1]Interatomic distances between the aromatic carbon atoms in the residues.

[2]X-ray crystallography of bovine α-La (1HFZ, Pike et al., 1996).

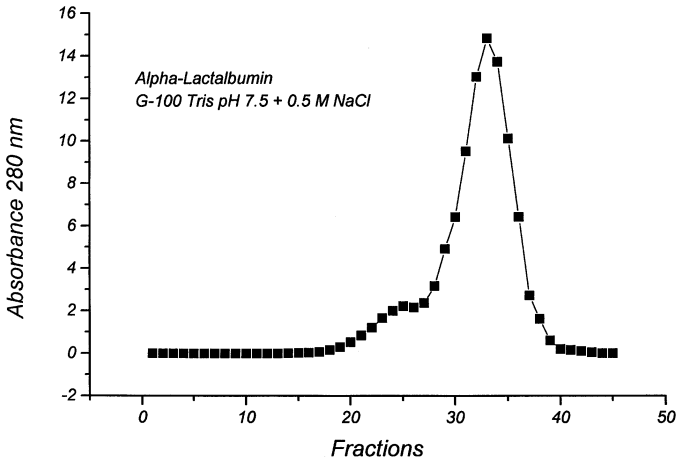

Figure 11. Sephadex G-100 chromatography of bovine α-lactalbumin after two cycles of DEAE cellulose columns with and without Ca^{2+}. Two fractions were found (I, 20 to 30 and II, 31 to 35). Both fractions were > 98% α-LA by SDS PAGE following reduction with 2-mercaptoethanol.

versely these experiments also demonstrate potential pitfalls in whey processing.

Implications of the "New View" and Molten Globules for Protein Structure

The following statements summarize some of the salient features of the new view and the molten globule state as discussed above for globular proteins in general and for α-LA in particular:

1. The majority of proteins contain some type of AFU, the nature of which depends upon primary structural elements.
2. During folding, the search of the possible protein energy landscape may be limited somewhat by primary structure, but the landscape certainly is much wider with many possible intermediates than previously thought.
3. Most likely many proteins can, under denaturing conditions, fall into a molten globule state. This state contains much native secondary structure, possibly some tertiary elements, but overall does not represent a compressed globular state. In addition, there is a freer rotation of aromatic side chains. As an aside, it is generally thought that chaperones may act to facilitate folding by binding to the molten globule state.
4. For α-LA and some other ligand binding proteins, removal of bound cations can lead to destabilization of structure, and perhaps lead to alternative states.
5. Internal hydration can lend some stability to the molten globule state, which by its open more flexible nature allows more surface for water interactions.
6. Misfolding and new surfaces for interactions can cause altered protein-protein interactions (aggregation) to issue forth from the molten globule state.

The question now arises, how can this "new view" be applied to casein structure?

Historic Views of Casein Structure

From all of the concepts regarding casein structure-function, which have been set forth over the years, two fundamental functions of casein can be envisioned: the effective transport of calcium and the self-associations that lead to the colloidal state (Farrell et al., 2002b; Holt and Sawyer, 1993; Swaisgood, 1982). Although casein has been studied for many years, the molecular structural basis for its function in self-association reactions in milk has been elusive. Historically, optical rotatory dispersion data from our laboratory demonstrated a lack of α-helix in the caseins, and since that was all that could be measured at the time, caseins were considered to be the model for random coil proteins (Farrell et al., 2002a). This would represent the "old view" of casein structure. However, as sequences became available, and CD was employed as tool for protein analysis (Creamer et al., 1981), the possibility of periodic structure was considered. Swaisgood (1982) was perhaps the first to suggest that the caseins were neither globular nor random coil proteins and that they could be composed of rather distinct functional domains. The next important step on the road to the understanding of casein structure may be the concept of Holt and Sawyer (1993), which suggests that caseins are rheomorphic in nature. In this instance, the "formed under flow" hypothesis suggests that casein structure is completely dynamic. In its extreme, this hypothesis may be considered as the "spaghetti plate" hypothesis, in that no regular structures occur until aggregates are formed. Supporting this hypothesis was the observation of Paulsson and Dejmek (1990) that pure caseins, when studied by DSC, showed flat endotherms on heating. These flat DSC scans are quite reminiscent of those found for α-LA either in the acid state (Figure 5) or in dilute solutions of guanidine as shown by Xie et al. (1991). It has thus been tempting to suggest that caseins in their native state exist in a type of molten globule state with a significant amount of open, but defined structure. The question remains whether or not this defined structure is localized to selective areas and persists during association reactions. Paulsson and Dejmek (1990) had indeed suggested this alternative view, that caseins exhibit no DSC peak, because they contain heat stable structures.

The idea that caseins have little or no fixed structure, however, remained as the prevalent view. In contrast,

data, which we collected in collaboration with Heino Susi, appeared to be at odds with this concept (Byler et al., 1988). We applied Susi's Raman methodologies, which he had developed for globular proteins, to the caseins. Analysis of the data revealed that all of the caseins had intricate Amide I profiles, similar to those obtained for globular proteins, with a moderate distribution of various secondary structures. The patterns of all purified caseins and their mixtures showed bands characteristic of reasonable amounts of β-turns and sheet and a modest amount of α-helix. It has, therefore, been somewhat difficult to reconcile the general view that structure gives rise to function with the random chain or rheomorphic hypothesis where function gives rise to structure.

Starting about 1990, we began to conduct a series of three-dimensional molecular modeling experiments on casein. In these studies, we attempted to derive casein structures from the basic Anfinsen hypothesis. The amount of conformational space sampled by the protein was constrained in the computer experiments to predicted secondary structure-derived from primary sequence data. We further constrained the global structure by requiring that it conform to Raman and FTIR limits. Finally, the number of turns was increased from algorithm predictions to correlate the spectroscopic data with the abundance of proline in the caseins. This was done because proline, while a structure-breaking residue for helix and sheet can be instrumental in the formation of turns in peptides and proteins (Kumosinski et al., 1993). Using these principles in conjunction with force-field calculations, we arrived at refined energy minimized working models for β-, κ-, and α_{s1}-CN (Farrell et al., 2002a). The κ-CN model is given in Figure 12 (Kumosinski et al., 1993).

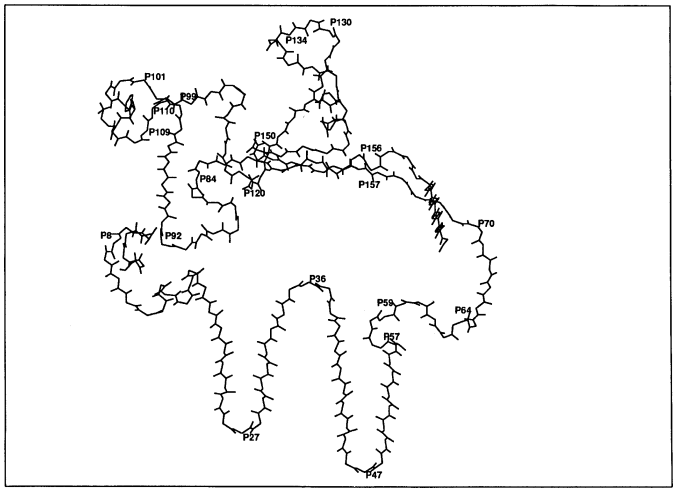

Figure 12. Energy minimized three-dimensional structure of κ-CN. The peptide backbone has the proline residues (P) labeled from Kumosinski et al. (1993).

Tensegrity: A "New View" of Casein Structure

Farrell et al. (2002a) proposed that the concepts of tensegrity may account for the protein structural interplay that occurs for caseins. Tensegrity is defined in architectural systems as a structure that stabilizes itself through a balance between compression and tension. Donald Ingber (1998) has pioneered the application of tensegrity to biological structures. A simple toy, shown in Figure 13, demonstrates how a few rigid struts joined by flexible elements can create a stable open structure with minimum building material. For the casein three-dimensional models, sheet-turn-sheet motifs centered on Pro residues provide the rigid rods, while loop and helix represent the more flexible elements (Farrell et al., 2002a). The κ-CN model (Figure 12) contains both the rigid rod-like structure (residues 14 to 64), as well as the flexible macropeptide region (105 to 169). The degree of flexibility of purified κ-CN is, however, limited by two factors: bound cations (Farrell et al., 1996) and a high degree of intermolecular disulfide bonding as first shown by Groves et al. (1992). However, isolated κ-CN exhibits a flat DSC profile

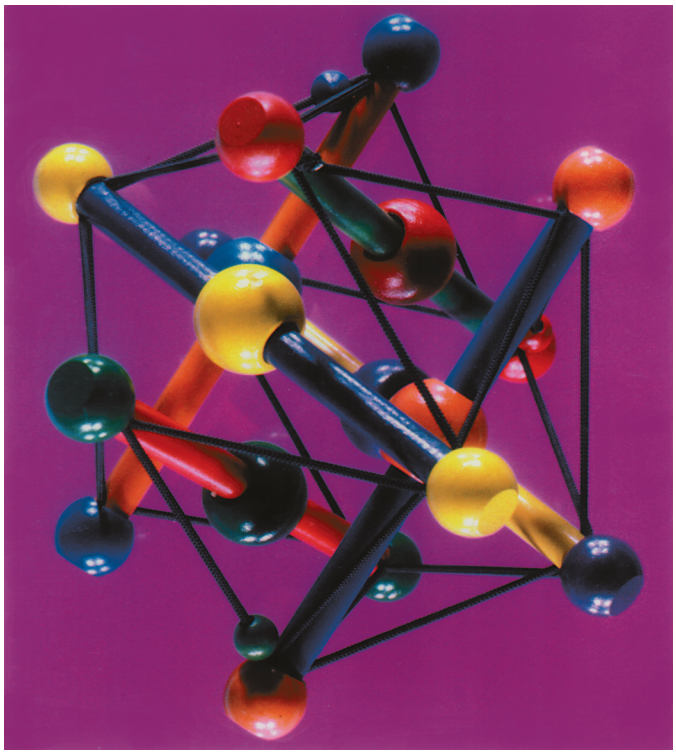

Figure 13. Tensegrity model, showing rigid rods and flexible connections. Courtesy of Manhattan Toys Inc.

Table 4. Comparison of secondary structural estimates for the κ-casein by two methods.

Method	Temperature, °C	% β-Sheet	% Turns	% Unspecified	% α-Helix
FTIR[1]	25	35 ± 3	25 ± 2	23 ± 4	17[2] ± 2
CD[3]	25	40 ± 1	26 ± 1	24 ± 1	9 ± 1
	10	36 ± 2	28 ± 1	24 ± 1	12 ± 1
	50	36 ± 2	27 ± 1	24 ± 1	14 ± 1
	70	36 ± 2	24 ± 1	20 ± 1	19 ± 1

[1]FTIR = Fourier transform infrared. Average of three determinations in PIPES-KCl aqueous at pH 6.75 (Farrell et al., 1996).

[2]For FTIR, includes 3_{10}-helix, bent strand and loop structures.

[3]For circular dichroism (CD), average fits, one determination (six accumulations at each temperature), the deviation for each element represents the error of the fit.

(Paulsson and Dejmek, 1990). Both FTIR and CD studies demonstrate significant amounts of secondary structure as shown in Table 4. In accordance with the DSC interpretations of Paulsson and Demjek (1990), these structures are quite stable and show no significant changes up to 70°C. These observations suggest that purified κ-CN may contain a heat-stable core with well-defined secondary structures. In fact, similar results have been reported by Graham et al. (1984) for β-CN, which also exhibits a flat DSC profile.

Alaimo et al. (1999) speculated that α_{s1}-CN may contain molten globule-like elements in the f136-196 peptide. Indeed, these heat-stable secondary structures, which are rich in hydrophobic residues, may be responsible for the ability of α_{s1}-CN to compete effectively with true molten globules for interaction sites on chaperones (Lin et al., 1995). Here, a true molten globule would be defined as a globular protein with native-like secondary structure, but unfolded tertiary structure (Arai and Kuwajima, 2000). It may then be suggested that α_{s1}-CN is a dead-end inhibitor of the chaperone-molten globule interaction, in that it contains secondary structural elements that bind to the chaperone, but it cannot form tertiary folds and so the complex does not dissociate. Caseins then may represent "native" molten globules with both persistent secondary structures (which maybe incapable of tertiary folds) and highly flexible elements as well. For casein, these persistent secondary structures such as the sheet-turn-sheet motifs may lead to self-association, but because of the compromise between tension and flexibility, no hydrophobic compression occurs and the proteins remain open and highly hydrated. In essence, they may move from secondary to quaternary structure without forming complex tertiary folded intermediates, which are characterized by DSC profiles for proteins like α-LA (Figure 5).

κ-CN: Restraints on Aggregation and/or Association by Metal Ions and Disulfide Bonds

For α-LA as noted above, Ca^{2+} plays a significant role in the stabilization of the protein. Removal of the ion causes changes in the DSC profile and may lead to altered denatured states. This can be seen in Table 1, where following MD simulations, the Rg of the apoprotein exceeds that of the holoprotein. The reverse is also true in that addition of Ca^{2+} enhances the folding reaction both in the simulations and in experimental studies. For purified κ-CN metal ions play a similar role. By electron microscopy (**EM**) purified κ-CN particles are somewhat spherical and have average radii of 9.0 nm. Treatment of κ-CN with EDTA leads to a broader size distribution (or radii) by dynamic light scattering, and an expanded shape as viewed by EM. The metal ion involvement was actually predicted by molecular modeling studies, which suggested the need for cations to minimize the model for a docked octamer. Removal of the metal ions also leads to formation of larger aggregates as viewed by sedimentation velocity studies (Farrell et al., 1999), and there is little accompanying change in secondary structure. It could be suggested that purified κ-CN actually represents a minimum energy form of an aggregate as defined by Jaenicke and Lilie (2000), because in vitro, the normal folding pathway would include interactions with other caseins. This fact, however, points to the potential uses of isolated fractions of casein as proteins with multipotential functionalities.

The open framework demonstrated by κ-CN aggregates (Farrell et al., 1996, 1999) most likely also accounts for its ability to stabilize casein micelles and form colloids with a variety of combinations of other caseins. An important unresolved issue is the timing of formation of κ-CN disulfide bonds in casein folding. Unless casein synthesis is synchronized, then the reaction of κ-CN with protein disulfide interchange enzyme (the latter occurs at the surface of the endoplasmic reticulum) could result in only intramolecular bonds. To yield its characteristic laddered forms (Groves et al., 1992), κ-CN molecules would have to be in queue. It has been speculated that the flexibility of the κ-CN allows it to concentrate on the surface of casein micelles perhaps by Ca^{2+} induced steering (Farrell et al., 1999,

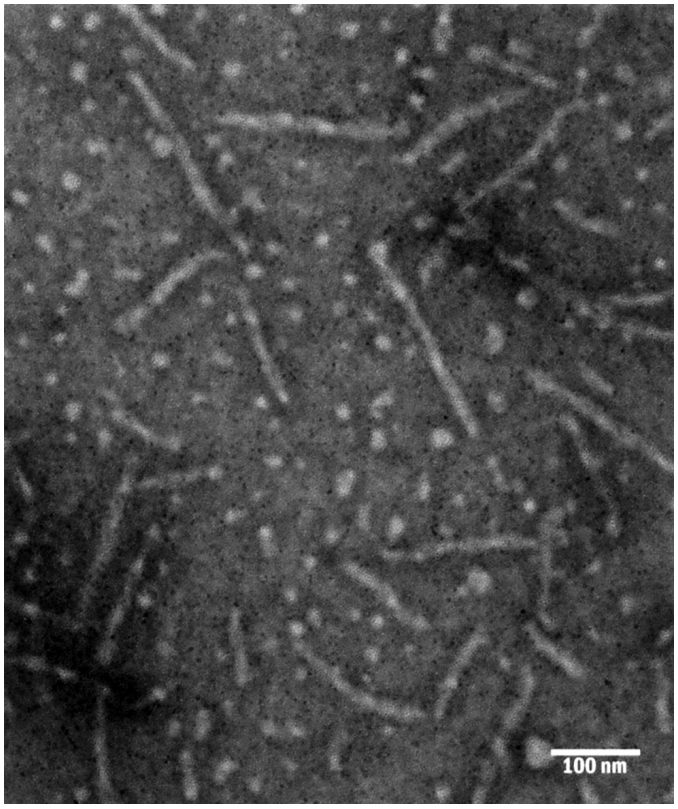

Figure 14. Transmission electron micrograph of a general field of negatively stained (uranyl acetate 2%). Reduced carboxymethylated-κ-casein showing both spherical particles and rod like shapes obtained at 37°C. Scale bar = 100 nm.

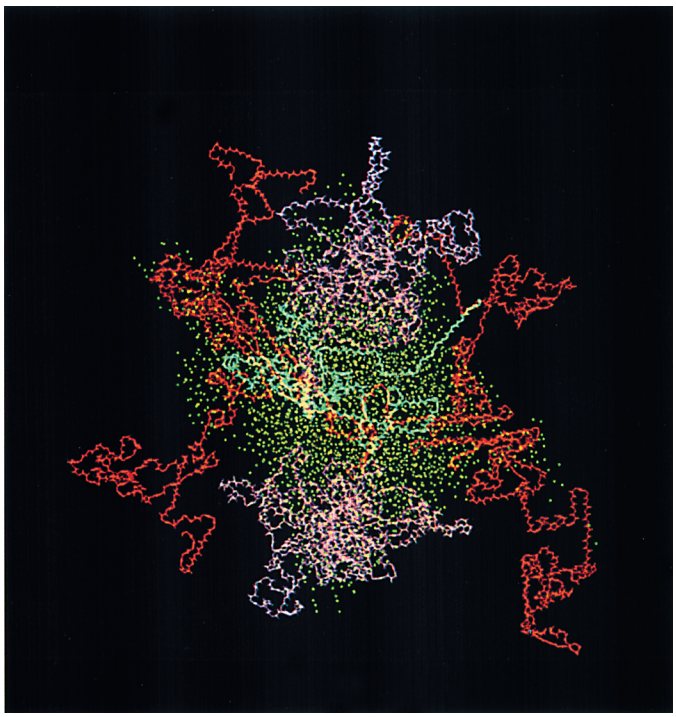

Figure 15. Backbone asymmetric structure of casein submicelle with water from droplet algorithm, 2823 water molecules; κ-CN in blue, α_{s1}-CN in red, β-CN magenta, and water molecules yellow green.

1996). This would occur after micelle formation in the later stages of secretion. This theory invokes a role for sulfhydryl oxidase (Swaisgood, 1982) in the formation of the κ-CN disulfides, perhaps even after milking.

To further study the role of disulfide bonding in κ-CN, the protein was reduced and carboxymethylated (**RCM**) as described in Groves et al. (1992). There was little change in the estimated amounts of secondary structure by FTIR or CD from that reported either for native or EDTA treated κ-CN (Table 4). However the potential for aggregation (or perhaps association) to a new form occurred. Just as the molten globule state of α-LA can yield new aggregates, κ-CN does as well in a very dramatic fashion. Negatively stained EM images of the κ-CN at 25°C showed particles much like those seen for the parent protein. The RCM protein, when heated to 37°C polymerized into long rod-like structures as seen in Figure 14. The sample also contains the spherical particles seen at 25°C for the native κ-CN (Farrell et al., 1996, 1999). Interestingly, our three-dimensional modeling studies actually predicted this form of aggregation (Kumonsinski et al., 1993) which could arise from stacking of β-sheet "legs" of the monomer of κ-CN (Figure 12). It is our contention that this potential for alternate forms of aggregation arises from the molten globule-like state of κ-CN. We further propose that this protein and α_{s1}-CN possess tensegrity structures, which enable them to pass directly from their molten globule-like states (which lack significant tertiary folds) to new aggregated forms or to associate and/or reassociate with each other during milk synthesis and secretion.

Casein Hydration

Finally one of the proposed features of the molten globule state of globular proteins is increased internal hydration due to the more open structure. For the caseins, the openness of the tensegrity structures and the three-dimensional models can readily accommodate an extremely high water content in the interior space. The water content of micelles and submicelles varies from 1 to 8 g of H_2O/g of protein (Farrell et al., 2002a; Mora-Gutierrez et al., 1997; Swaisgood, 1982). In a series of studies on casein water interactions, Mora Gutierrez et al. (1997) have used ^{17}O NMR to probe and enumerate sources of bound, trapped, and preferentially absorbed water molecules. These cavities and voids have been correlated with the three-dimensional model of sodium

caseinate (Farrell et al., 2002a; Mora-Gutierrez et al., 1997), which is shown, partially hydrated, in Figure 15. Thus, the tensegrity hypothesis coupled with the possibility of a static molten globule state can account for the physical and chemical properties of caseins, while explaining how the molecules can be both rigid and flexible at the same time and exhibit highly hydrated backbones.

CONCLUSION

In conclusion, as the field of structural biology evolves, we will learn more about the basic structures of the milk proteins. Perhaps two-dimensional X-ray crystallography or EM reconstruction may yield higher resolution models of the caseins, possibly even at the atomic level. However, as the field progresses, we need to be aware of the opportunities these advances present for enhancement of dairy products or for adaptation of milk proteins as food ingredients. To more fully take advantage of the unique properties of the milk proteins, we will need new alternative and cost effective means of fractionation. In addition, new insights into present processing methodologies will be required to yield the innovative processes needed for future applications. So, we are at the beginning of an exciting era in protein chemistry and hopefully milk technology.

REFERENCES

Akers, R. M., T. B. McFadden, W. E. Beal, A. J. Guidry, and H. M. Farrell. 1986. Radioimmunoassay for measurement of bovine α-lactalbumin in serum, milk and tissue culture media. J. Dairy Res. 53:419–429.

Alaimo, M. H., H. M. Farrell, Jr., and M. W. Germann. 1999. Conformational analysis of the hydrophobic peptide α_{s1}-casein(136-196). Biochim. Biophys. Acta 1431:410–420.

Arai, M., and K. Kuwajima. 2000. Role of the molten globule state in protein folding. Adv. Protein Chem. 53:209–282.

Byler, D. M., H. M., Farrell, Jr., and H. Susi. 1988. Raman spectroscopic study of casein structure. J. Dairy Sci. 71:2622–2629.

Creamer, L. K., T. Richardson, and D. A. D. Parry. 1981. Secondary structure of α_{s1}- and β-caseins in solution. Arch. Biochem. Biophys. 211:689–696.

Dolgikh, D. A., L. V. Abaturov, I. A. Bolotina, E. V. Brazhnikov, V. E. Bychkova, R. I. Glimanshin, Y. O. Lebedev, G. V. Semisotnov, E. I. Tiktopulo, and O. B. Ptitsyn. 1985. Compact state of a protein molecule with pronounced small-scale mobility. Eur. Biophys. J. 13:109–121.

Farrell, H. M., Jr., E. M. Brown, P. D. Hoagland, and E. L. Malin. 2002a. Higher order structures of caseins: a paradox. In Advanced Dairy Chemistry. P. Fox and P. McSweeney, eds. (in press). Kluwer Academic, NY, NY.

Farrell, H. M., Jr., T. F. Kumosinski, P. H. Cooke, P. D. Hoagland, E. D. Wickham, J. J. Unruh, and M. L. Groves. 1999. Environmental influences on the particle sizes of purified κ-casein: Metal effect. Int. Dairy J. 9:193–199.

Farrell, H. M., Jr., T. F. Kumosinski, P. H. Cooke, G. King, P. D. Hoagland, E. D. Wickham, H. J. Dower, and M. L. Groves. 1996. Particle sizes of purified κ-casein: metal effect. J. Protein Chem. 15:435–445.

Farrell, H. M., Jr., T. F. Kumosinski, E. L. Malin, and E. M. Brown. 2002b. The caseins of milk as calcium binding proteins. Pages 97–134 in Calcium Binding Protein-Protocols. H. J. Vogel, ed. Humana Press, Totowa, NJ.

Graham, E. R. B., G. M. Malcolm, and H. A. McKenzie. 1984. On the conformation of β-casein. Int. J. Biol. Macromol. 6:155–161.

Groves, M. L., H. J. Dower, and H. M. Farrell, Jr. 1992. Reexamination of the polymeric distribution of κ-casein isolated from bovine milk. J. Protein Chem. 11:21–28.

Håkansson, A., M. Svensson, A.-K. Mossberg, H. Sabharwal, S. Linse, I. Lazou, B. Lönnerdel, and C. Svanborg. 2000. A folding variant of α-lactalbumin with bactericidal activity against *Streptococcus pneumoniae*. Macromol. Microbiol. 38:589–600.

Holt, C., and L. Sawyer. 1993. Caseins as rheomorphic proteins. J. Chem. Soc. Faraday 89:2683–2690.

Ingber, D. E. 1998. The architecture of life. Sci. Am. Jan. pp 48–57.

Jaenicke, R., and H. Lilie. 2000. Folding and association of oligomeric and multimeric proteins. Adv. Protein Chem. 53:329–401.

Kataoka, M., K. Kuwajima, F. Tokonaga, and Y. Goto. 1997. Structural characterization of the molten globule of α-lactalbumin by solution X-ray scattering. Protein Sci. 6:422–430.

Kumosinski, T. F., H. Pessen, and H. M. Farrell, Jr. 1982. Structure and mechanism of action of riboflavin binding protein. Arch. Biochem. Biophys. 214:714–725.

Kumosinski, T. F., E. M. Brown, and H. M. Farrell, Jr. 1993. Three dimensional molecular modeling of caseins: an energy minimized κ-casein structure. J. Dairy Sci. 76:2507–2520.

Kumonsinski, T. F., and J. J. Unruh. 1996. Quantitation of the global secondary structure of globular proteins by FTIR spectroscopy. Talanta 43:199–219.

Lin, Z., F. P. Schwarz, and E. Eisenstein. 1995. The hydrophobic nature of Gro EL-substrate binding. J. Biol. Chem. 270:1011–1014.

Mora-Gutierrez, A., T. F. Kumosinski, and H. M. Farrell, Jr. 1997. Oxygen-17 NMR studies of bovine and caprine caseins: hydration and activity in deuterated sugar solutions. J. Agric. Food Chem. 12:4545–4553.

Onuchic, J. N., H. Nymeyer, A. E. Garcia, J. Chahine, and N. D. Socci. 2000. The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. Adv. Protein Chem. 53:88–153.

Paulsson, M., and P. Dejmek. 1990. Thermal denaturation of whey proteins in mixtures with caseins studied by calorimetry. J. Dairy Sci. 73:590–600.

Peng, Z.-Y., and L. C. Wu. 2000. Autonomous protein folding units. Adv. Protein Chem. 53:1–49.

Pike, A. C., K. Brew, and K. R. Achorya. 1996. Crystal structures of guinea-pig, goat and bovine α-lactalbumin. Structure 4:691–703.

Saito, M. 1999. Molecular dynamics model structures for the molten globule state of α-lactalbumin. Protein Eng. 12:1097–1104.

Smith, L. J., C. M. Dobson, and W. F. van Gunsteren. 1999. Side chain conformational disorder in a molten globule. J. Mol. Biol. 286:1567–1580.

Svensson, M., H. Sabbarwal, A. Håkansson, A.-K. Mossberg, P. Lipniunas, H. Leffler, C. Svanborg, and S. Linse. 1999. Molecular characterization of α-lactalbumin folding variants that induce apoptosis in tumor cells. J. Biol. Chem. 274:6388–6396.

Swaisgood, H. E. 1982. Chemistry of milk proteins. Pages 1–60 *in* Developments in Dairy Chemistry. P. F. Fox ed. Applied Science London, United Kingdom.

Xie, D., V. Bhakuni, and E. Freire. 1991. Calorimetric determination of the energetics of the molten globule intermediate in protein folding. Biochemistry 30:10673–10678.